Supplemental Information

Judson et al., "*In Vitro* Screening of Environmental Chemicals for Targeted Testing Prioritization – The ToxCast Project"

Data Availability: All of the data described in this article is available from the EPA ToxCast web site: http://www.epa.gov/ncct/toxcast. The complete sets of AC50/LEC values are contained in a series of 9 files packaged into a zip files with an accompanying README describing the contents. Additionally, the large supplemental Excel tables referred to below are also on the ToxCast web site.

Chemical Library: A complete listing of the 320 chemical library and a structure data file (SDF) are available for download at http://www.epa.gov/NCCT/dsstox/sdf_toxcst.html. The dataset includes 309 unique chemical structures, along with three sets of triplicate samples (same source/lot/batch) and 5 sets of duplicate samples (obtained from different commercial sources) used for assessing assay replicability. The majority of these chemicals are current or former food-use pesticide active ingredients designed to be bioactive, whereas the rest are industrial chemicals of environmental relevance. Chemical information was quality reviewed and structure-annotated within the DSSTox project (for more information on DSSTox quality review procedures and standard chemical fields, see: http://www.epa.gov/ncct/dsstox). Chemicals comprising the ToxCast Phase-I library were procured and plated by BioFocus DPI (South San Francisco, CA) under contract to EPA. Supplier-provided certificates of analysis indicated purity >97% for the large majority of chemicals (87%), and >90% purity for all but a few instances of technical grade or known mixtures. Follow-up analysis of a chemical solution plate identical to those used in the majority of assays was performed using LC/MS (liquid chromatography mass spectrometry), and confirmed mass identification, stability, and purity for over 83% of the chemical library. The analysis methods employed were inadequate for confirming presence of some chemicals in the library; stability studies are underway for some of the remaining chemicals. A QC

summary result mapped to chemical sample will be provided on the ToxCast website (http://www.epa.gov/ncct/toxcast/).

AC50/LEC Calculation: For assay sets where saturating behavior was expected, concentration-response curves were fit to a 3 or 4 parameter Hill functions and an AC50 (half maximal activity concentration)(Inglese et al. 2006) value was derived. For assay sets where saturating behavior was not typically seen (such as in RNA expression assays) a Lowest Effective Concentration (LEC) was identified as the lowest concentration at which there was a statistically significant change from the concurrent negative control. LEC values were typically at lower concentrations than AC50 values because, even in the case of saturating behavior, the LEC will occur before the AC50 and can be determined even if 50% activity is never reached. Assay sets for which the characteristic concentration is an LEC tended to have more active chemicals per assay and to have lower characteristic concentrations on average. For cell-based assays, cytotoxicity can cause artifactual appearance of specific cellular phenotypes. Therefore, specific chemical-assay combinations were called inactive if the AC50/LEC was equal to or greater than the lowest cytotoxic concentration measured in the same cell line.

Cell-free HTS Assays: A collection of biochemical assays measuring binding constants and enzyme inhibition values. There are a total of 239 endpoints. Chemicals were initially screened at a single concentration in duplicate wells at a concentration of 10 μM for CYP450 assays and 25 μM for all others. Chemicals-assay pairs that showed significant activity were then run in concentration response format, from which an AC50 value was extracted. For concentration response, 8 concentrations were tested in the ranges 0.00914-20 μM for CYP assays and 0.0229-50 μM for non-CYP assays. These assays were run by Caliper Life Sciences (Hanover, MD) under contract to EPA. Short assay descriptions are available at http://www.caliperls.com/products/contract-research/in-vitro/. Values for all endpoints were imported into GeneData Screener (Basel, Switzerland) and normalized to solvent controls from the same plate as the tested compounds and expressed as percent of neutral control. Concentration-response curves

were fit in GeneData Screener (Basel, Switzerland) Condeseo module using the Hill equation and AC50 values were determined. Data were fit to 4 parameter Hill model as first choice; secondarily a 3 parameter Hill fit was used, generally with a fixed top of curve at 100% of positive control. Detailed description of the methods, data and analysis methods for this dataset are described in a separate publication (Knudsen et al. 2009).

Cell-based HTS Assays: These assays measuring binding constants and enzyme inhibition values, primarily for nuclear receptors, but also including an assay for p53. The targets include AR, ER, FXR, LXR, GR, PPARα, PPARδ, PPARγ, RXRα, RXRβ, PXR and TRβ (Inglese et al. 2007; Wilkinson et al. 2008). Each of the nuclear receptor targets was measured in both agonist and antagonist mode, but most of the antagonist data is still undergoing analysis due to possible interference between cytotoxicity and assay activity, so antagonist mode data were not used in the present analysis. There were a total of 18 endpoints. Assays were run at the NIH Chemical Genomics Center (Rockville, MD) as part of the Tox21 collaboration(Collins et al. 2008). Normalized data from the NCGC were imported into GeneData Screener (Basel, Switzerland) Condeseo Module and fit to 4 parameter Hill model as first choice; secondarily a 3 parameter Hill fit was used, generally with a fixed top of curve at 100% of control unless good indication of partial agonism. Minimum 25% of control efficacy was required to consider the chemical to be active, and the R^2 for the curve fit was required to be ≥ 0.5. Values where only the highest concentration exceeded 50% activity were excluded to eliminate weak or false positives. Detailed description of the methods, data and analysis methods for this dataset will be presented in a separate publication.

High Content Cell-Imaging (Vernetti et al. 2008): This is 2 sets of high-content imaging assays that test a chemical's effects on a range of phenotypes in either the human hepatoma cell line HepG2, or rat primary hepatocytes. Although partially deficient in some important liver phase I and phase II enzymes, HepG2 can metabolize compounds resulting in production of toxic metabolites or inactivation of a toxic parent molecule similar to what occurs in the intact liver. Freshly cultured rat primary hepatocytes have a more complete metabolic capacity, relative to the HepG2. The HepG2 data set included

11 unique phenotypic endpoints: cell cycle status, cell loss, nuclear size, oxidative stress, stress kinase, p53 activation, mitotic arrest, mitochondrial membrane potential, mitochondrial mass, microtubule function and microtubule destabilization. The rat primary hepatocyte panel included 8 measurements: cell count, nuclear size, apoptosis, lysosomal mass, cell count, nuclear size, mitochondrial membrane potential/mass, and hepatic steatosis. Measurements were made at three time points in 384 well microplates using a Cellomics ArrayScan VTi and the appropriate BioApplications. Values were imported into GeneData Screener, normalized to solvent controls and expressed as percent of neutral control. Concentration-response curves were fit in the GeneData Condeseo module using the Hill equation and AC50 values determined. If an AC50 was calculated for Cell Loss, it was used to filter other endpoints measured at the same exposure time. If the AC50 value was greater then the Cell Loss AC50, the AC50 of the endpoint was set to 200 μM (the highest concentration tested). Assays were run by Cellumen, Inc. (Pittsburgh, PA) under contract to EPA. Detailed description of the methods, data and analysis methods for this dataset will be presented in a separate publication.

Gene Expression in Human Primary Hepatocytes: This is a collection of multiplexed gene expression assays focused on Phase I and II metabolic enzymes and transporters. Concentration- and time-response profiles of chemicals are measured by changes in 1) the expression of key nuclear receptor target genes, 2) CYP1A enzymatic activity (EROD), and 3) cell morphology. Fourteen gene targets were monitored by quantitative nuclease protection assay: six representative cytochrome P450 genes, four hepatic transporters, three Phase II conjugating enzymes, and one endogenous metabolism gene involved in cholesterol synthesis. Cells were exposed at 5 concentrations (0.004-40 μM) for 6, 24 or 48 hr. Gene targets are sentinels for five nuclear receptor signaling pathways: AhR, CAR, PXR, FXR, and PPARα. All gene are associated with nuclear receptor pathways: CYP1A1 and CYP1A2 with AHR; ABCB1, ABCG2, CYP2B6, CYP2C9, CYP2C19 and UGT1A1 with CAR; CYP3A4, GSTA2, SLCO1B1 and SULT2A1 with PXR; HMGCS2 with PPARA; and ABCB11 with FXR. Assays were run in primary human hepatocyte cultures (LeCluyse et al. 2005) by CellzDirect Invitrogen Inc.

(Durham, NC), in collaboration with EPA. Detailed description of the methods, data and analysis methods for this dataset will be presented in a separate publication.

Multiplex Transcription Reporter Assays: A multiple reporter transcription unit (MRTU) library consisting of 48 transcription factor binding sites was transfected into the HepG2 human liver hepatoma cell line as previously described (Romanov et al. 2008). In addition to the cis-acting reporter genes (CIS), a modification of the approach was used to generate a trans-system (TRANS) with a mammalian one-hybrid assay consisting of an additional 25 MRTU library reporting the activity of nuclear receptor super-family members. Based on an initial cytotoxicity screen, the maximum tolerated concentration (MTC) was derived as one-third the calculated IC50 or, if no IC50 was determined, the MTC was set to 100 μM. Chemicals were then tested in the CIS and TRANS assays at seven concentrations starting at the MTC and followed by three-fold serial dilutions. These assays were performed by Attagene Inc. (Morrisville NC) under contract to EPA. Detailed description of the methods, data and analysis methods for this dataset will be presented in a separate publication.

Biologically Multiplexed Activity Profiling (BioMAP): These assays use human cell systems to characterize effects relevant to human tissue and inflammatory disease biology following exposure to chemicals. These are complex cell systems, using primary human cells, leverage cellular regulatory networks to detect and distinguish chemicals with a broad range of target mechanisms and biological processes relevant to human toxicity. Chemicals were assayed at four concentrations in eight BioMAP cell systems, with a total of 87 assay endpoints. Assays were performed by BioSeek, Inc. (South San Francisco CA) (Berg et al. 2006). Detailed description of the methods, data and analysis methods for this dataset are given in a separate publication(Houck et al. 2009).

Phase I & II XME cytotoxicity assays: This is a set of microarray based cytotoxicity assays that measure the effect of adding Phase I and II human xenobiotic metabolizing enzyme (XME) and transporter activities. There are a total of 4 cytotoxicity endpoints, with either no XME, Phase I only, Phase II only, or Phase I and II XME added to the

assay process. Cytotoxicity assays were run in Hep3B cells at 9 concentrations in the range 0.0146-960 μM and AC50 values calculated (Lee et al. 2008). These assays were run by Solidus Inc. (Troy, NY) in collaboration with EPA and L'Oreal. Detailed description of the methods, data and analysis methods for this dataset will be presented in a separate publication.

High-throughput genotoxicity screening: The GreenScreen high content *GADD45a-GFP* high-throughput screening (HTS) genotoxicity assay was measured in concentration response format (Hastwell et al. 2006). The assay was run in TK6 cells at 3 concentrations in the range 12.5-50, or 50-200 μM depending on cytotoxicity. This assay was run by Gentronix, Inc. (Manchester, UK) in collaboration with EPA. Detailed description of the methods, data and analysis methods for this dataset are given in a separate publication(Knight et al. 2009).

Real-Time Cell Electronic Sensing (RT-CES): The RT-CES system was used to measure the time-dependent response to chemicals. The RT-CES system utilizes specially designed microtiter plates embedded with electronic cell sensor arrays in the bottom of each well to measure minute changes in cell number, morphology and adhesion via direct measurement of time-dependent cell impedance (Xing et al. 2006; Xing et al. 2005; Solly et al. 2004). Assays were run in A549 cells at 8 concentrations in the range 0.047-100 μM. All screening was carried out by ACEA Biosciences (San Diego, CA) under contract to EPA. For each chemical 7 parameters were derived from time-dependent impedance curves. The first parameter is the cell growth AC50 values derived from fitting the concentration-dependent values of the average growth rate to a Hill equation. Values of AC50 above the highest concentration tested (100 μM) are discarded. The remaining parameters are Lowest Observed Concentration (LOC) values, which are the LOC at which specific curve shapes were first seen. Detailed description of the methods, data and analysis methods for this dataset will be presented in a separate publication.

Data Reproducibility: One measure of data quality is the rate of replication of hits in the assays. Within the 320 substance tested, there were 3 sets of triplicates from the same lot and batch of chemical. For each of the assay technology sets, we calculated an overall concordance and hit concordance rate using the triplicates. For each pair in the triplicate (comparing A with B, B with C and A with C) we ask if the chemicals are both hits (meaning they have an AC50/LEC at or below the highest concentration tested, and below the concentration at which cytotoxicity was seen for cell-based assays), both non-hits or if they disagree. We get total concordance if all 3 comparisons register hits or all 3 of them are non-hits, and non-concordance otherwise. The overall concordance rate is given by A+D/(A+B+C+D) where A is the number of cases (a pair of chemicals and one assay) where both chemicals are a hit, C is the number of cases where the first is a hit and the second is a non-hit, C is the number of cases where the first is a non-hit and the second is a hit and D is the number of cases where both are non-hits. The hit concordance is equal to A/(A+B+C). Supplemental Table 5 gives the overall and hit concordance for 7 of the 9 assay technologies. The cell-based HTS assays were run without replicates and the high throughput genotoxicity assay had no hits among any of the replicated chemicals.

From Supplemental Table 5, we see that the overall concordance rates are at least 90% except for the cell-imaging assays, meaning that if one sample is a hit or one a non-hit, the other will agree most of the time. The hit concordance is lower for most of the cell-based assays than for the cell-free biochemical assays, indicating that it is more often difficult to make definitive calls of hits in these assays. One important use *in vitro* assays is as quantitative input variables to statistical models of *in vivo* toxicity. If the numerical AC50/LEC values are not reliably reproducible, assays are not useful in this context. The individual BioMAP assays do not have good hit concordance, as seen from the table, but we have shown elsewhere that the quantitative pattern of activity for a chemical across all of the assays (the BioMAP profile) is stable and reproducible(Houck et al. 2009). A similar approach is being assessed with the high-content cell imaging assays. It will be important to assess whether lack of hit concordance arises more from false positive or false negatives. The quantitative nuclease protection assays show poor hit concordance at

the earliest time point of 6 hours, which reflects the fact that this is too early for cells to have reliably activated the transcription machinery as a result of chemical exposure.

Univariate Statistical Analysis: The first step in predictive model building was univariate feature selection. First, *in vivo* endpoints were dichotomized, with LEL values of greater than 2,000 mg/kg/day being set to inactive. Only the 309 unique chemicals were included in the calculations. Two statistical tests were performed: 1) a t-test comparing continuous *in vitro* results, and 2) a Fisher's exact test comparing dichotomized *in vitro* results wherein AC50/LEC values >100 were considered inactive. We discarded all associations where, in the continuous case, the mean of the *in vivo* positive chemical incidence was less than that of the *in vivo* negative incidence. Such associations could be interpreted as "protective" effects of the chemical, but most of these instances arose when there were very few chemicals positive for the *in vivo* endpoint, so they are statistically suspect. We also discarded all associations in which there were fewer than 2 chemicals for which there was both an assay and an endpoint hit. For selected endpoints, we repeated the univariate association analysis using standard permutation testing procedure to help correct for multiple comparisons. As a first pass criterion, *in vivo* endpoint data was initially permuted 100 times. If the calculated relative risk value or Fisher p-value fell within the upper 20^{th} percentile, then the permutation test was carried out for an additional 100,000 permutations. If the calculated relative risk and Fisher p-value both fell within the upper 95^{th} percentile of the second permutation, then the corresponding assay was selected as a feature to be used in the predictive models. The univariate predictors from the initial test against rat liver tumors still showed significance at the p=0.05 level.

Relationship between ToxCast Assays and Human Disease: The significance of these pathway hits can be put into context by using the mapping from pathways to disease recently published by Gohlke et al. (Gohlke et al. 2009). That study developed associations between published pathways (KEGG) and diseases, and environmental factors using information on gene / disease links from genetic variation data extracted from the Genetic Association database (GAD) (Becker et al. 2004) and the Comparative

Toxicogenomics Database (CTD) (Davis et al. 2009; Mattingly et al. 2006). Supplemental Table 4 lists all of the KEGG pathways for which we have assays, the number of chemicals hitting the related pathway assay and the number of diseases enriched for those pathways for CTD and GAD (from Table 2 of Golke et al.). The 15 KEGG pathways most highly enriched for disease links from CTD and GAD are listed. We have at least one chemical hitting 10 of15 GAD top pathways and 9 of 15 of the CTD top pathways. Because the HTS data includes multiple assays that allow us to probe a majority of these pathways that are significantly related to human diseases, the activity of ToxCast chemicals in these pathways can be an *in vitro* starting point for understanding potential human toxicities associated with these chemicals.

Supplemental Tables and Figures

Supplemental Table 1: Complete listing of all assays.

File name: Table_S2_ToxCastAssayMaster.xls available from the ToxCast web site *http://www.epa.gov/ncct/toxcast*

- SOURCE_NAME_AID - Short code for the assay
- ASSAY_NAME - Name of the assay
- ASSAY_SOURCE - Company or organization providing the assay data
- ASSAY_DESCRIPTION - Descriptive information for the assay
- ASSAY_CODE - Short assay code
- ASSAY_COMPONENTS – Codes for each assay measurement, used as the header in the data files
- ASSAY_UNITS - units
- ASSAY_CATALOG_NUMBER - Vendor catalog number (For Caliper / Novascreen only)
- ASSAY_URL -URL to the assay provider
- ASSAY_CATEGORY1,2 - In vitro (cell or cell-free), computational, in vivo
- ASSAY_PHENOTYPE - Usually, the broad class of in vivo toxicity for in vivo toxicity data
- SPECIES - Species for the gene, protein or cells
- ASSAY_TARGET - Gene that is target of the assay, usually = GENE_SYMBOL
- ASSAY_TARGET_FAMILY - Family of target, such as kinase, GPCR
- ASSAY_TARGET_SOURCE - Tissue source of the target protein
- ASSAY_TARGET_SOURCE_TYPE - Type of source such as cell line or tissue
- ASSAY_GENE_ID - Entrez Gene ID
- GENE_NAME - Entrez gene symbol
- ASSAY_TECHNOLOGY - Type of technology used
- ASSAY_MODE - Assay Mode (e.g. agonist, antagonist, fold-change)

- ASSAY_POSITIVE_CONTROL - Positive control chemical
- ASSAY_SOLVENT_CONTROL - Solvent control compound
- ASSAY_REFERENCE_COMPOUND - Positive control compound
- ASSAY_NOTE - Other notes
- ASSAY_SUBSTRATE_NAME - Name of the substrate
- ASSAY_SUBSTRATE_CONCENTRATION_M - Concentration of the substrate for the assay
- ASSAY_ENZYME_SUBSTRATE_AFFINITY_KM_M - Molar substrate affinity for the enzyme
- ASSAY_ENZYME_VMAX - V(max) for enzyme assays
- ASSAY_REACTION - Description of reaction taking place in assay
- ASSAY_ATP_CONCENTRATION_M - ATP concentration
- ASSAY_ENZYME_AFFINITY_ATP_KM_M – KM for ATP affinity
- ASSAY_LIGAND_NAME - Name of ligand for displacement assays
- ASSAY_LIGAND CONCENTRATION_M - Ligand concentration
- ASSAY_KD_M - KD for the assay
- ASSAY_BMAX - B(max) for the assay

Supplemental Table 2 List of all assay-chemical combinations with an AC50/LEC value equal to or below 1 μM. The table provides the chemical, the assay code from Table S1, the gene associated with the assay where available and the AC50/LEC value in units of μM.

File name: Table_S2_Hits_under_1uM.xls available from the ToxCast web site *http://www.epa.gov/ncct/toxcast*

Supplemental Table 3: List of all pathway-chemical LEC values. All values are in μM and chemical-assay pairs with no activity are assigned a default value of 1000000 μM.

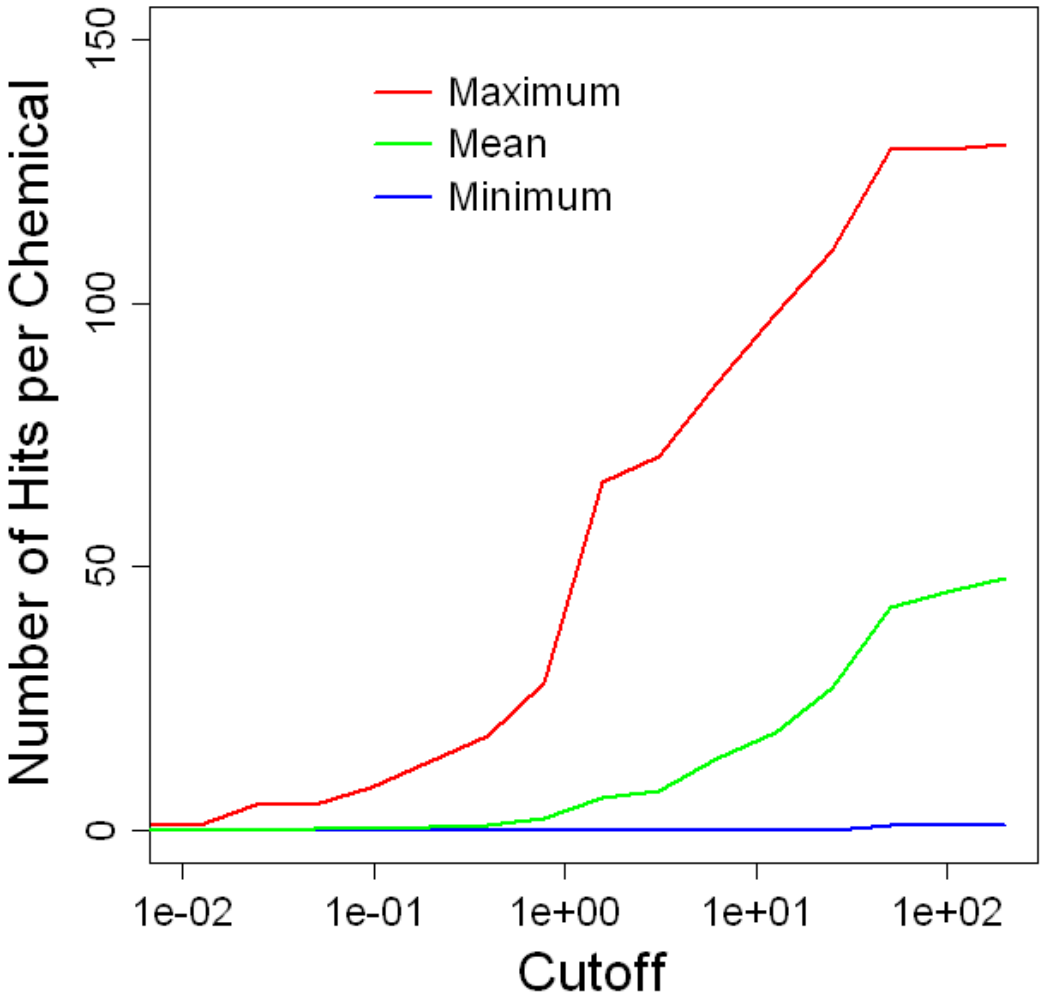
File name: Table_S3_S4_Pathway_assay_results.xls, Sheet labeled (Table S3) available from the ToxCast web site http://www.epa.gov/ncct/toxcast

Supplemental Table 4: Correlation between number of chemical hits in ToxCast and diseases and environmental factor hits from Gohlke et al. (Gohlke et al. 2009) for the KEGG pathways.

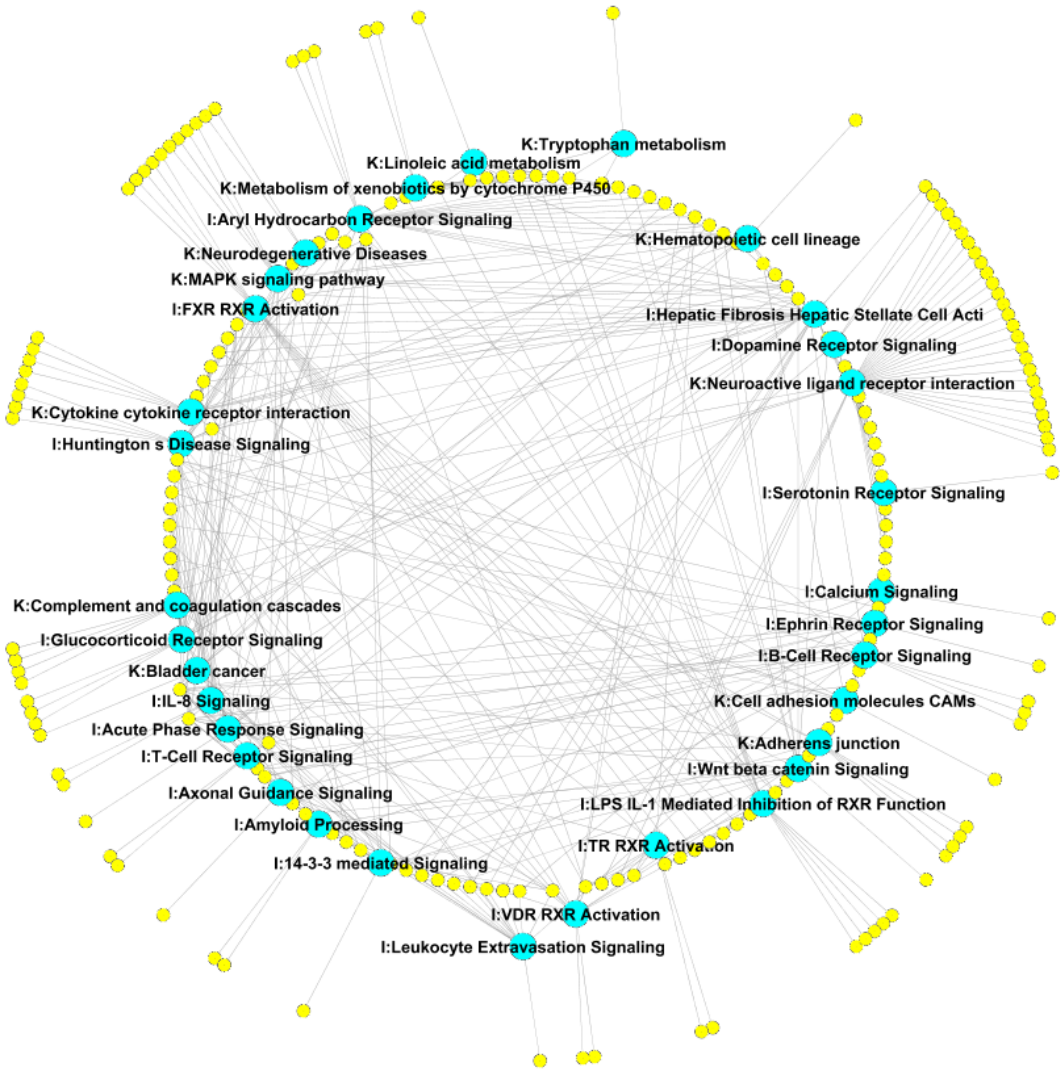
File name: Table_S3_S4_Pathway_ assay_results.xls, Sheet labeled (Table S4) available from the ToxCast web site http://www.epa.gov/ncct/toxcast

Assay Set	Overall Concordance	Hit Concordance
Cell-free HTS	99	85
Multiplex Transcription Reporter	99	87
Biologically Multiplexed Activity Profiling (BioMAP)	90	32
Phase I & II XME Cytotoxicity	96	84
Real Time Cell Electronic Sensing	90	62
High Content Cell-Imaging (HepG2) 1 Hr	96	50
High Content Cell-Imaging (HepG2) 24 Hr	86	44
High Content Cell-Imaging (HepG2) 72 Hr	86	33
High Content Cell-Imaging (RPH) 1 Hr	94	60
High Content Cell-Imaging (RPH) 24 Hr	86	23
High Content Cell-Imaging (RPH) 48 Hr	83	37
Quantitative Nuclease Protection 6 Hr	86	36
Quantitative Nuclease Protection 24 Hr	90	66
Quantitative Nuclease Protection 48 Hr	92	80

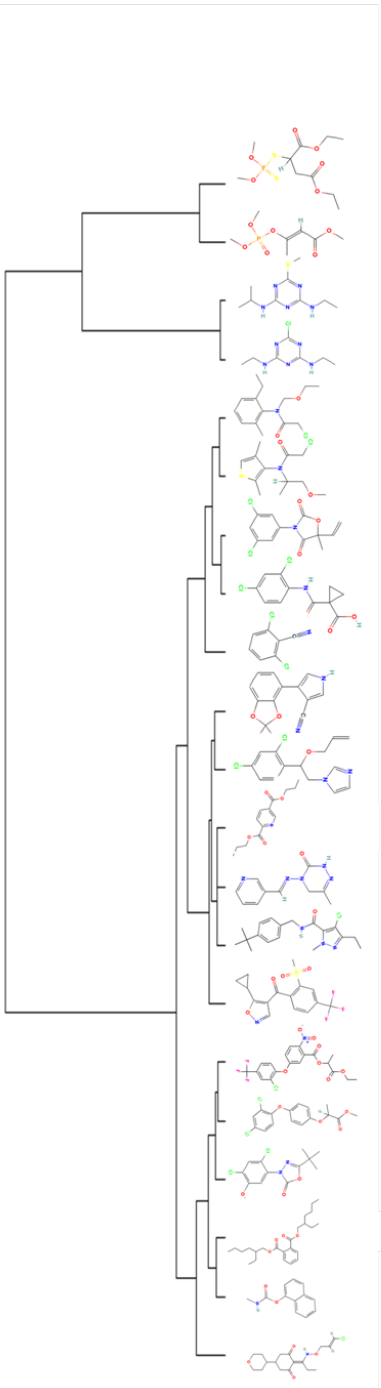
Supplemental Table 5: Replication statistics. Reproducibility of the assay data was evaluated using 3 sets of chemicals run in triplicate. For each assay technology, the results were compared for each pair of chemicals in the triplicate for each assay. In each case, the 2 chemicals could both be hits, both non-hits, or one hit and one non-hit. The overall concordance rate is given by A+D/(A+B+C+D) where A is the number of cases where both chemicals are a hit, C is the number of cases where the first is a hit and the second is a non-hit, C is the number of cases where the first is a non-hit and the second is a hit and D is the number of cases where both are non-hits. The hit concordance is equal to A/(A+B+C). For the high content cell imaging and quantitative nuclease protection assays, results are reported by time point. For the former, we also break out results by the two types of cells used: HepG2 and rat primary hepatocyte (RPH).



Supplemental Figure 1: Maximum, minimum and mean number of hits per chemical as a function of the threshold or cutoff used to define activity. The cutoff is in units of μM. Maximum is the maximum number of assays that any chemical is active for, where a hit is defined as having an AC50/LEC equal to or below the cutoff. Minimum is defined likewise. Mean is the mean number of hits across all of the chemicals for a given cutoff.



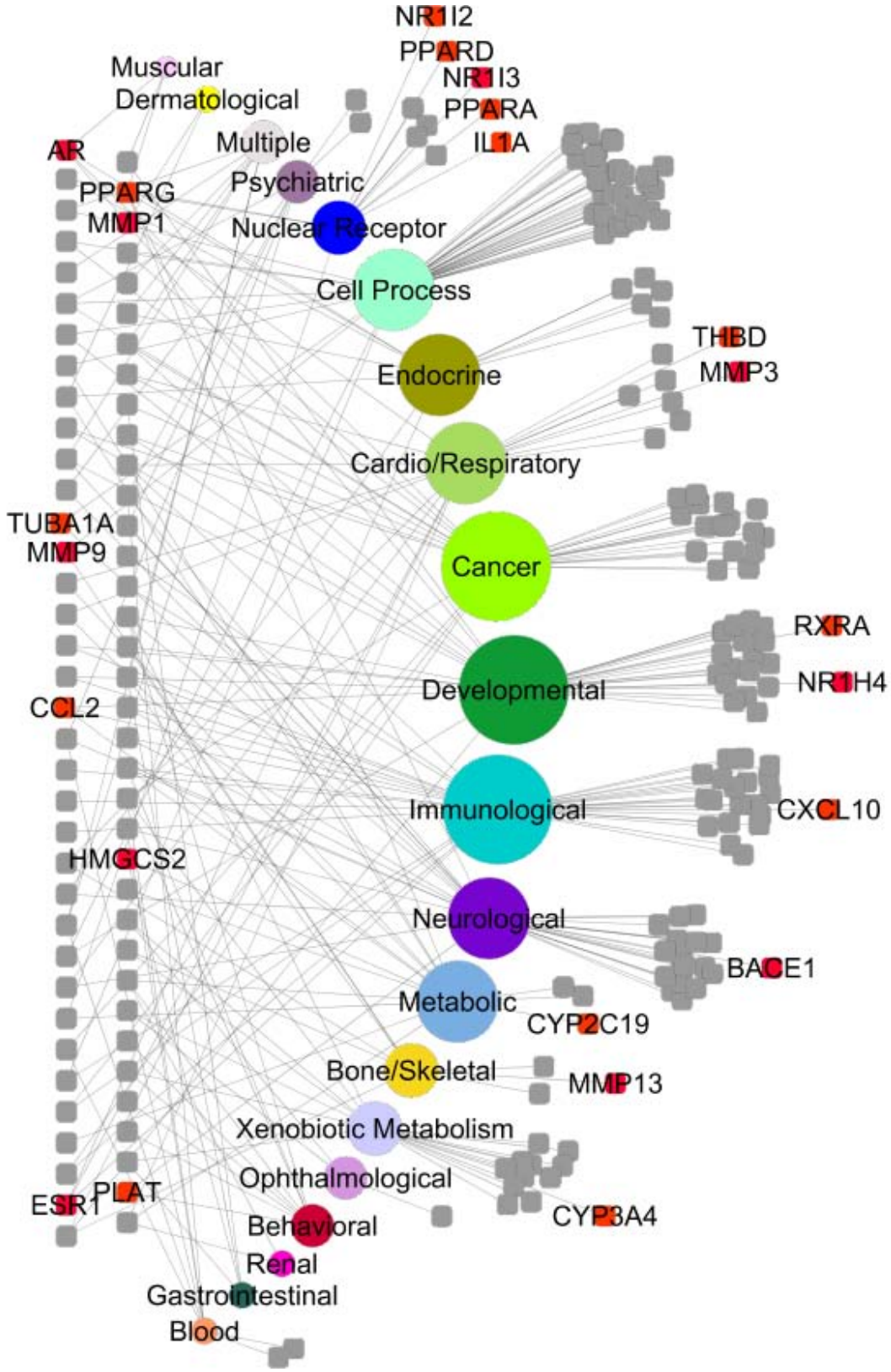
Supplemental Figure 2: Diagram of the network of minimal human pathways and genes for which assays are available in the current dataset. Pathways are shown in blue and labeled. Associated genes are shown in yellow. An important point of this diagram is that even these minimally overlapping pathways share many key genes. “K”: KEGG, “I”: Ingenuity pathways.



	AR	PPARA	PPARG	HMGCS2	CCL2
Malathion	■		■		
Mevinphos			■		■
Ametryn			■	■	
Simazine			■		■
Acetochlor	■		■	■	
Dimethenamid	■		■		
Vinclozolin	■		■		
Cyclanilide			■	■	
Dichlobenil					
Fludioxonil			■		
Imazalil		■	■		
2,5-Pyridinedicarboxylic acid, dipropyl ester			■		■
Pymetrozine					
Tebufenpyrad			■	■	
Isoxaflutole			■		■
Lactofen	■	■	■	■	
Diclofop-methyl	■	■	■		■
Oxadiazon	■		■		■
Diethylhexyl phthalate		■	■		
Carbaryl			■		
Tepraloxydim					

Supplemental Figure 3: The 21 unique chemicals that caused rat liver tumors were clustered using Tanimoto similarity scoring from the PubChem database's structure clustering tool. These compounds were associated with 5 *in vitro* assays (Assay codes refer to Table S1) that were selected as features for predicting rat liver tumors. The

symbol for the gene being measured is also indicated. This figure demonstrates the difficulty in using a structure based approach to identify biological activity related to carcinogenesis.


NR1I2
PPARD
NR1I3
PPARA
IL1A
Muscular
Dermatological
Multiple
Psychiatric
Nuclear Receptor
AR
PPARG
MMP1
Cell Process
THBD
MMP3
Endocrine
Cardio/Respiratory
TUBA1A
MMP9
Cancer
RXRA
NR1H4
Developmental
CCL2
Immunological
CXCL10
HMGCS2
Neurological
BACE1
Metabolic
CYP2C19
Bone/Skeletal
MMP13
Xenobiotic Metabolism
Ophthalmological
ESR1
PLAT
CYP3A4
Behavioral
Renal
Gastrointestinal
Blood

Supplemental Figure 4: Network diagram of the genetic loci and corresponding disorder and biological process classes probed by ToxCast Phase I HTS Assays. Colors denote grouped disease types defined by OMIM disease-gene associations and based partially on previously defined disease classes (Goh et al. 2007). Node size corresponds to the number of distinct loci present in each disorder class. Genetic loci are illustrated in gray, with the exception of those discussed in the text, which are indicated in red.

References for Supplemental Material

Inglese J, Auld DS, Jadhav A, Johnson RL, Simeonov A, Yasgar A, et al. 2006. Quantitative high-throughput screening: a titration-based approach that efficiently identifies biological activities in large chemical libraries. Proceedings of the National Academy of Sciences of the United States of America 103(31): 11473-11478.

Knudsen T, Houck K, Judson RS, Singh A, Mortensen H, Reif D, et al. 2009. Biochemical Activities of 320 ToxCast Chemicals Evaluated Across 239 Functional Targets. Nature Chem Bio submitted.

Inglese J, Johnson RL, Simeonov A, Xia M, Zheng W, Austin CP, et al. 2007. High-throughput screening assays for the identification of chemical probes. Nature chemical biology 3(8): 466-479.

Wilkinson J, Hayes S, Thompson D, Whitney P, Bi K. 2008. Compound profiling using a panel of steroid hormone receptor cell-based assays. J Biomol Screen 13(8): 755-765.

Collins FS, Gray GM, Bucher JR. 2008. Toxicology. Transforming environmental health protection. Science 319(5865): 906-907.

Vernetti L, Irwin W, Giuliano KA, Gough A, Johnston P, Taylor DL. 2008. Cellular Systems Biology Applied to Pre-Clinical Safety Testing: A Case Study of CellCiphr Cytotoxicity Profiling. In: Drug Efficacy, Safety and Biologics Discovery: Emerging Technologies and Tools (Xu JJ, Ekins S, eds).

LeCluyse EL, Alexandre E, Hamilton GA, Viollon-Abadie C, Coon DJ, Jolley S, et al. 2005. Isolation and culture of primary human hepatocytes. Methods Mol Biol 290: 207-229.

Romanov S, Medvedev A, Gambarian M, Poltoratskaya N, Moeser M, Medvedeva L, et al. 2008. Homogeneous reporter system enables quantitative functional assessment of multiple transcription factors. Nat Methods 5(3): 253-260.

Berg EL, Kunkel EJ, Hytopoulos E, Plavec I. 2006. Characterization of compound mechanisms and secondary activities by BioMAP analysis. Journal of pharmacological and toxicological methods 53(1): 67-74.

Houck KA, Dix DJ, Judson RS, Kavlock RJ, Yang J, Berg EL. 2009. Profiling bioactivity of the ToxCast chemical library using BioMAP primary human cell systems. J Biomol Screen 14(9): 1054-1066.

Lee MY, Kumar RA, Sukumaran SM, Hogg MG, Clark DS, Dordick JS. 2008. Three-dimensional cellular microarray for high-throughput toxicology assays. Proceedings of the National Academy of Sciences of the United States of America 105(1): 59-63.

Hastwell PW, Chai LL, Roberts KJ, Webster TW, Harvey JS, Rees RW, et al. 2006. High-specificity and high-sensitivity genotoxicity assessment in a human cell line: validation of the GreenScreen HC GADD45a-GFP genotoxicity assay. Mutat Res 607(2): 160-175.

Knight AW, Little S, Houck K, Dix D, Judson R, Richard A, et al. 2009. Evaluation of high-throughput genotoxicity assays used in profiling the US EPA ToxCast chemicals. Regul Toxicol Pharmacol 55(2): 188-199.

Xing JZ, Zhu L, Gabos S, Xie L. 2006. Microelectronic cell sensor assay for detection of cytotoxicity and prediction of acute toxicity. Toxicol In Vitro 20(6): 995-1004.

Xing JZ, Zhu L, Jackson JA, Gabos S, Sun XJ, Wang XB, et al. 2005. Dynamic monitoring of cytotoxicity on microelectronic sensors. Chem Res Toxicol 18(2): 154-161.

Solly K, Wang X, Xu X, Strulovici B, Zheng W. 2004. Application of real-time cell electronic sensing (RT-CES) technology to cell-based assays. Assay Drug Dev Technol 2(4): 363-372.

Gohlke JM, Thomas R, Zhang Y, Rosenstein MC, Davis AP, Murphy C, et al. 2009. Genetic and environmental pathways to complex diseases. BMC Syst Biol 3: 46.

Becker KG, Barnes KC, Bright TJ, Wang SA. 2004. The genetic association database. Nature genetics 36(5): 431-432.

Davis AP, Murphy CG, Saraceni-Richards CA, Rosenstein MC, Wiegers TC, Mattingly CJ. 2009. Comparative Toxicogenomics Database: a knowledgebase and discovery tool for chemical-gene-disease networks. Nucleic Acids Res 37(Database issue): D786-792.

Mattingly CJ, Rosenstein MC, Davis AP, Colby GT, Forrest JN, Jr., Boyer JL. 2006. The comparative toxicogenomics database: a cross-species resource for building chemical-gene interaction networks. Toxicol Sci 92(2): 587-595.

Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabasi AL. 2007. The human disease network. Proceedings of the National Academy of Sciences of the United States of America 104(21): 8685-8690.